AD-A111 181 NAVAL BIODYNAMICS LAB NEW ORLEANS LA F/G 5/9
TASK ANALYSIS AND THE ABILITY REQUIREMENTS OF TASKS: COLLECTED --ETC(U)
SEP 81 R H SHANNON, R C CARTER
UNCLASSIFIED NBDL-81R009 NL

1 OF 1
AD
A111181

END
DATE
FILMED
3-82
DTIC

30 LEVEL II

NBDL - 81R009

Task Analysis and the Ability Requirements of Tasks: Collected Papers

R. H. Shannon and R. C. Carter

September 1981

NAVAL BIODYNAMICS LABORATORY
New Orleans, Louisiana

82

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NBDL-81R009	2. GOVT ACCESSION NO. AD-A111 181	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Task Analysis and the Ability Requirements of Tasks: Collected Papers		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER NBDL-81R009
7. AUTHOR(s) R. H. Shannon and R. C. Carter		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Biodynamics Laboratory New Orleans, Louisiana 70189		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Project F58524 Task Area ZF5852406 Work Unit MF58.524-00205027
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research & Development Command Bethesda, MD 20014		12. REPORT DATE September 1981
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Task Analysis, Human Performance

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This is a collection of papers on the topic of using task analysis to identify the abilities required for jobs. Identification of ability requirements is a first step in the process of assembling performance tests that are related to a particular job. Once assembled such tests could be used for many purposes. One use would be selection of personnel who are especially gifted with the required attributes. The use to which the authors' efforts are directed is environmental research. Environmental research is the study of human performance in unusual environments (e.g., vibration, ship motion,

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) impact) in order to determine whether and to what degree the environment affects performance. Such research is more relevant to the Navy when the human performance studied is related to Navy jobs.

NBDL - 81R009

Task Analysis and the Ability Requirements of Tasks: Collected Papers

R. H. Shannon and R. C. Carter

September 1981

Bureau of Medicine and Surgery
Work Unit No. MF58.524-002=5027

Approved by

Channing L. Ewing, M. D.
Chief Scientist

Released by

Captain J. E. Wenger MC USN
Commanding Officer

Naval Biodynamics Laboratory
Box 29407
New Orleans, LA 70189

SUMMARY PAGE

PROBLEM

Human performance tests are used for a variety of research and personnel management purposes. In each case, the tests are supposed to measure some ability which is needed for performance of a job. The problem is that there is no well-established method for specifying the ability requirements of jobs, and hence it is difficult to choose appropriate performance tests.

FINDINGS

Task analysis can be used to specify the ability requirements of jobs. A variety of task analysis methods are illustrated by the papers in this collection, and their application to enumeration of ability requirements is described.

RECOMMENDATION

It is recommended that task analysis be employed prior to performance test selection so that tests can be chosen to be relevant to Navy jobs.

The work was funded by the Naval Medical Research and Development Command and by the Biological Sciences Division of the Office of Naval Research.

Table of Contents

Proceedings of the 24th Annual Meeting of the Human Factors Society, Los Angeles CA
13 - 17 October 1980

TASK ANALYTIC APPROACH TO HUMAN PERFORMANCE BATTERY DEVELOPMENT

Richard H. Shannon
Naval Biodynamics Laboratory, New Orleans, LA 70189

ABSTRACT

Task analytic methods were used to isolate critical components of naval student flight performance. This goal was accomplished by utilizing factor and regression analyses to compare student maneuver errors during primary training to the overall phase grades of primary, basic and advanced. The results indicated that flight ability skills appear to be best measured by basic transitions, coordination flying and entries to dirty configurations. These items could be measured within a flight simulator. In a laboratory, these skills could best be measured by a performance battery which contained memory, problem-solving, continuous tracking tasks, and a time-sharing capability between continuous tracking and discrete tasks. In conclusion, the task analytic procedure was determined to be a feasible and useful undertaking in the development of performance measurement systems such as the PETER project.

INTRODUCTION

The Naval Biodynamics Laboratory (NBL) in New Orleans is presently attempting to develop a human performance battery (PETER, Performance Evaluation Tests for Environmental Research), which will be used to study behavior under unusual and adverse conditions. To facilitate the accomplishment of this goal, a task analysis of various U. S. Navy jobs and work stations will be conducted. In general, the steps to be followed from inventory development to battery implementation will be: (1) Research and evaluate the existing literature in the areas of task analysis, battery development, and performance assessment; (2) Isolate the critical work stations onboard U. S. Navy ships which are important to its peacetime and combat missions; (3) Complete a task inventory of these critical work stations; (4) Develop a questionnaire in order to isolate the recurring critical elements pertaining to a particular work station; (5) Translate these elements into perceptual-psychomotor behavioral components; and (6) Design a perceptual-psychomotor battery to be used at NBL which is compatible with the existing work of PETER.

Previously, the author had followed a similar outline for the purpose of developing a performance battery which would assess student pilot behavior. Although the battery was never developed or implemented, most of the necessary work had been conducted. The purpose of this report will be to discuss this prior effort because the methodology, data collection, and analyses will have a direct impact upon the future direction of the PETER program. In another paper, the author (Shannon, 1980) had discussed the development of maneuver task inventories, the isolation of critical inflight student errors, the validation and utility of the methodology and taxonomic classification model, and finally the potential impact of this effort upon future NBL research. Therefore, the major emphasis of this investigation will be upon the predictive validity of the model, the translation of tasks into behavioral components, and the ultimate design of a performance battery.

METHOD

The taxonomic model, which was discussed in the previous paragraph, was developed in order to classify recurring student pilot flight errors during primary flight training in the T-34 aircraft. This taxonomy depicts pilot performance as a joint function of continuous and discrete communications, scanning, and controlling operations that occur in a four dimensional inflight environment of pitch, roll, yaw, and thrust. These tasks are performed as part of an information processing loop between the pilot, aircraft, and the environment. This feedback loop contains the components of display, sensory, cognition, motor and control. This classification model, which can be found in Shannon (1980), uses two operations, nine activities, and 35 functional objectives. Discrete operations are defined as individually distinct movements or mediating responses elicited by environmental cues. Continuous operations contain those tasks involving multidimensional tracking responses to either contact cues outside the cockpit or flight instrument cues within the cockpit. A task activity is a qualitative category whose main function involves either a sensory, cognition, motor, or coordinated perceptual-motor task. The nine activities used in this paper are: continuous tasks involving control of the pitch axis, roll axis, yaw axis, thrust axis and the brakes; and discrete tasks involving control of aircraft subsystems (procedural), judgment and planning (anticipation/planning), transfer of information (communication), and searching and scanning (monitor). Finally, functional objectives are tasks having the same activity, goal-orientation, sensory cues and task elements. For example, continuous pitch axis control contains the objectives to maintain altitude, airspeed, nose attitude, and stick pressure.

The data in this study were collected prior to 1977 and the revision of the naval aviation training program. At that time, training consisted of four phases: primary, basic, advanced and the replacement air group. This study will concern itself with the first three phases. A content analysis was performed on the grading sheets at the primary phase of two samples of 39 students, a total sample size of 78. These grading sheets contained data on both sides pertaining to a student's performance on a particular hop. On the front page, the maneuvers that were performed with their associated grade (abovaverage to unsatisfactory) are listed. When a below average or unsatisfactory grade was assigned, the instructor had to justify his maneuver grade by listing on the back page of the grading sheet the errors committed by that student. The analyses in this paper were performed on these written comments and classified using the taxonomic model described in the previous paragraph. There were 2,161 errors for 550 below average/ unsatisfactory grades in the first sample, while the second sample committed 2,093 errors for 529 below average/ unsatisfactory grades on the 15 maneuver and three global (headwork, procedures, basic airwork) items which demonstrated deficient performance.

In addition, these errors were further classified as to fundamental flight attitude for continuous operations and task purpose for discrete operations. The fundamental flight attitudes used in the analysis of continuous tasks were straight climb, climbing turn, straight descent, descending turn, level turn, straight and level, and ground. On the basis of these seven attitudes, it was possible to characterize a maneuver according to transitions from one flight attitude to another or the maintenance of a particular flight attitude. The comparison of discrete tasks was performed using the purpose for task execution as the basis for the analysis. For example, the following reasons for performance of the functional objective "anticipation of a particular position" were: to initiate turn completion, to commence a turn, to communicate a voice report, and to add/retract power. This effort resulted in 89 and 74 categories of continuous and discrete flight errors, respectively.

These two sets of error categories were factor analyzed separately by the principal axis and rotated by the varimax. These methods tend to maximize the percent of common variance shared by a factor while minimizing the overlap between factors. In addition, the factor analytic output resulted in factor scores which are the sum of the products between maneuver weights and error category standard scores on each maneuver. All error categories with a factor score greater than 1.0 and a maneuver factor loading greater than .21 were used in the prediction model. Student error scores were recomputed from this more selective item pool by multiplying the errors in each category by the maneuver factor loading. Each of these products were then summed within a maneuver and across maneuvers to give a recomputed factor error score for each student. This is not the first time factor analysis has been used with respect to naval aviation training (Bair, Lockman, Martoccia, 1956; Booth, Berkshire, 1968; Wherry, Waters, 1960; Bale, Smith, Ambler, 1972). In each of these cases, the analysis used grades during the various phases of training. The present study differed, however, by using instructor comments translated into errors as the basis for the analysis.

Since content and concurrent validity of the primary phase data had been established by another study (Shannon, 1980), the next step was to establish predictive validity and the relationship of initial flight performance to later phases of training. A series of stepwise regression analyses were performed using the overall phase grades for primary, basic and advanced as criterion measures on the first sample of 39 students. Multiple correlations were determined using the recomputed student error scores on each factor as independent variables. Cross-validation consisted of using the same maneuver loadings, error categories, factor structure and regression beta weights on the three criterion measure equations from the first sample to analyze the second sample. With the data validated, both samples were combined (N=78) and a third series of regression equations were computed using the same phase grades.

RESULTS AND DISCUSSION

From the two separate varimax rotations, twelve factors were identified, five continuous and seven discrete, each explaining 39% and 64% of the total variance respectively. Table 1 contains the results of these two rotations with only the factor loadings of .21 and above being shown. These factors were called continuous tasks during field entries, emergency landings, unusual attitudes, spins and basic airwork performance; and discrete tasks that occurred during transitions, stall recovery -nose high or nose low, emergency landings, clearing/ break turns, slow flight, and course rule performance. Table 2 outlines each factor by name and its significant functional activities, functional objectives, and fundamental flight attitudes (continuous tasks) or flight purpose (discrete tasks). In other words, the significant error categories are listed by factor.

Table 3 contains a correlational matrix for both predictor factor variables and criterion stage grade variables. The correlations in the matrix indicate that five of the twelve factors have a rather consistent affect on performance throughout pilot training. The significant continuous factors indicate that maintaining altitude and airspeed (pitch), maintaining distance (roll) and maintaining heading (yaw) in various fundamental flight attitudes (turns, descents, climbs, straight and level) are important. On the other hand, the critical discrete factors are: (1) to use the throttle and to anticipate airspeed and altitude so as to initiate a descent or to level-off from a descent; (2) to raise the gear on departure, to anticipate altitude for level-off, and to determine geographic location so as to adhere to inflight procedural rules; and (3) to use throttle, prop, flaps appropriately when entering slow flight. In addition, two factors, field entries (continuous) and transitions (discrete), are highly correlated (r = .70). The important grading items outlined by these five factors are - headwork, basic airwork, procedures, standard and home field entries, slow flight, transitions and course rules. In general, these results indicate that basic airwork skills, physical coordination, good scan pattern, knowledge of procedures, ability to plan ahead of the aircraft, capability to time-share and handle workload stress are important attributes throughout training. These skills or abilities seem to be best measured during basic transitions, coordination flying, steep turns, and entries to dirty configuration (lowering the landing gear). These items could be studied either in a flight simulator or under actual flight conditions. In a laboratory setting, these skills could best be measured by a battery which contained memory, problem-solving, and continuous tracking tasks as well as a time-sharing capability between continuous tracking and discrete tasks.

Table 3 also contains multiple correlations of the factor analytic scores at the three phases of training. Three sets of data are presented: first sample of 39, second sample of 39 using the same beta weights for validation purposes, and the combined sample. The following conclusions can be drawn from this table: (1) the taxonomic classification of student error has predictive validity, (2) performance in primary training does predict future airborne behavior, which may indicate that there is a general factor of "flight ability" and that this factor concurrently exists with specific criterion shifts and changes within the task (changing-task model), (3) there is a decrease in the predictive capability of the data base the further a student progresses in training, which may demonstrate that there are specific as well as a hierarchy of skills being developed throughout training (changing-subject model). These last two statements direct attention to the controversy surrounding temporal changes in ability-skill relationships (Alvares, Hulin, 1972).

The changing-task model (Fleishman, 1966) specifies that, in order to explain the decreasing relationships between abilities measured early in training and performance measures obtained later, one must relate different abilities to performance at the various stages of skill acquisition. In other words, the interval measuring system of flight instructors within naval aviation changes over time in that their emphasis upon specific critical components or skills shifts between training stages but remains relatively constant within a stage. They expect more precise behavior and better integrated performance as a student progresses through training. Fleishman (1966) stated that the importance of nonmotor abilities (such as, sensory, verbal) decreases systematically with practice, relative to motor abilities. This idea can further be supported by Table 3 which illustrates that continuous motor factors throughout training appear to be more predictive than the more cognitive discrete tasks. In other words, the sensory, cognitive and motor capabilities of the pilot are important, but the total integration and measurement of these abilities can only be attained by observing the output variables of motor capacity as a part of a feedback loop.

The changing-subject model (Adams, 1957; Humphreys, 1960) assumes that living organisms show constant change, and that measures taken at successive intervals depict this change. In other words, student abilities can be seen as systematically changing over time within a constant task structure. For example, Underwood (1977) claimed that three of his laboratory tasks (free recall, list differentiation, interference susceptibility) measured different memory abilities because of the low correlations between these tests. However, when the scores during the stable period of these three tests (Days 7-14) were factor analyzed, they demonstrated a high amount of explained common variance (63 percent). This difference appeared to be related to the level of training of the subjects. Underwood studied one day of learning while the factor analyzed material investigated the period after six days.

To summarize, both models appear to explain the decreasing correlations of primary phase errors with later phases of training. The process of flight instruction and assessment can be viewed as a feedback loop between the student, and the flight instructor. The instructor establishes an internal measurement system of proficient and critical performance based upon student behavior and conversations with his peers. On the other hand, the student conforms to the desires and criticisms of the flight instructor. His performance becomes more integrated as specific abilities-skills are intertwined within a hierarchy as components of a more advanced ability-skill. Within this context, the task and subject both change over time, and therefore should better explain why correlations decrease as the distance between stages of training increase.

The implications of the results of this paper are significant to the future of the PETER project. The author concludes from these analyses that the task analytic procedure, of isolating critical components of a job and then developing performance measurement systems, training packages and human engineering applications based upon these components, is a feasible undertaking. Therefore, task analysis will be an integral part in the future development of a battery to assess performance within unusual environments at the Naval Biodynamics Laboratory.

REFERENCES

Adams, J. A., The Relationship Between Certain Measures of Ability and the Acquisition of a Psychomotor Response. Journal of General Psychology, 1957, 56, 121-134.

Alvares, K., Hulin, C., Two Explanations of Temporal Changes in Ability-Skill Relationships: A Literature Review and Theroetical Analysis. Human Factors, 1972, 14(4), 295-308.

Bair, J. T., Lockman, F. F., and Martoccia, C. T., Validity and Factor Analyses of Naval Air Training Predictor and Criterion Measures. J. Appl. Psychol. 4:213-219, 1956.

Bale, R. M., Smith, M. J., and Ambler, R. K., Factor Analysis of Undergraduate and Post-graduate Flight Training Grades. Aerospace Med. 43(4):372-375, 1972.

Booth, R. F., and Berkshire, J. R., Factor Analysis of Aviation Training Measures and Post Training Performance Evaluations. NAMI-1050, Pensacola, FL: Naval Aerospace Medical Institute, 1968.

Fleishman, E. A., Human Abilities and the Acquisition of Skill. In E. A. Bilodeau (ed.) Acquisition of Skill, New York: Academic Press, 1966, 147-167.

Humphreys, L. G., Investigations of the Simplex. Psychometrika, 1960, 4, 313-323.

Shannon, R., The Validity of Task Analytic Information to Human Performance Research in Unusual Environments, Proceedings Human Factors Society, 1980.

Underwood, B., Boruch, R., Malmi, R., The Composition of Episodic Memory, Evanston, Illinois: Northwestern Uni, 1977.

Wherry, R. J., Jr., and Waters, L. K., Factor Analysis of Primary and Basic Stages of Flight Training: Advance Jet Pipeline Students. NSAM-254, Pensacola, FL: Naval School of Aviation Medicine, 1960.

TABLE 1: TWO VARIMAX ROTATIONS DEPICTING FACTOR LOADINGS BY GRADED ITEM, FACTOR AND OPERATION

	CONTINUOUS OPERATIONS					DISCRETE OPERATIONS						
GRADED FLIGHT ITEMS	I	II	III	IV	V	I	II	III	IV	V	VI	VII
HEADWORK (HD)	.45	.64			.40			.26		.27		.68
BASIC AIRWORK (BAW)	.55	.27			.54	.69				.28		
PROCEDURES (PROC)		.61								.53		.24
APPROACH/LANDING/TAKEOFF (APP)	.23	.22			.21				.41			
STANDARD FIELD ENTRY (SFE)	.83					.59			.32	.32		
HOME FIELD ENTRY (HFE)	.88					.78			.24			
EMERGENCY LANDING PRACTICE (ELP)		.67						.26	.56	.32		
HIGH ALTITUDE EMERGENCY (HAE)		.83						.77				
LOW ALTITUDE EMERGENCY (LAE)		.65						.81				
SLOW FLIGHT (SF)					.58	.23				.85		
TRANSITION (TRAN)					.59	.92						
POWER-OFF STALL (POS)	.23	.24							.78			
APPROACH TURN STALL (ATS)				.27			.99					
PRECISION SPIN (PS)				.61					.94			
ACCIDENTAL SPIN (ACC SP)				.63			.24				.95	
UNUSUAL ATTITUDES (UA)			.68				.87				.50	
WINGOVER (WO)			.66									
COURSE RULES (CR)												.44
PERCENT COMMON VARIANCE	10.8	13.2	4.	4.4	6.11	3.7	10.5	8.1	12.4	7.9	6.7	4.5

TABLE 2: SIGNIFICANT ERROR CATEGORIES BY FACTOR AND OPERATION

FACTOR NAME (MANEUVER):	FUNCTIONAL OBJECTIVE (FUNCTIONAL ACTIVITY)	FUNDAMENTAL FLIGHT ATTITUDE/FLIGHT PURPOSE
A. CONTINUOUS TASK DURING:		
1. FIELD ENTRIES (SFE, HFE, BAW)	MAINTAIN ALTITUDE/AIRSPEED (PITCH) AND DISTANCE (ROLL)	STRAIGHT DESCENT, LEVEL TURN, STRAIGHT AND LEVEL
2. EMERGENCY LANDINGS (HD, PROC, ELP, HAE, LAE)	MAINTAIN AIRSPEED (PITCH) AND AND RATE OF DESCENT/DISTANCE (ROLL)	DESCENDING TURN
3. UNUSUAL ATTITUDES (UA, WO)	MAINTAIN A CONTINUOUS RATE OF NOSE ATTITUDE (PITCH) AND ANGLE OF BANK (ROLL) WHILE MAINTAINING BALANCED FLIGHT (YAW)	CLIMBING AND DESCENDING TURN

TABLE 2: (CONTINUED)

4. SPINS (PS, ACC SP)	MAINTAIN A SPECIFIC NOSE ATTITUDE (PITCH) ON ENTRY; OR NEUTRAL STICK PRESSURE (PITCH), NEUTRAL RUDDER (YAW) AND WINGS LEVEL (ROLL) ON RECOVERY	STRAIGHT DESCENT ON ENTRY AND RECOVERY
5. BASIC AIRWORK PERFORMANCE (BAW, SF, TRAN)	MAINTAIN ALTITUDE/AIRSPEED (PITCH) AND HEADING (YAW)	STRAIGHT AND LEVEL, DESCENDING TURN, STRAIGHT CLIMB
B. DISCRETE TASKS DURING:		
1. TRANSITIONS (BAW, SFE, HFE, TRAN)	USE THROTTLE (PROCEDURES), AND ANTICIPATE AIRSPEED/ALTITUDE	TO LEVEL-OFF FROM A DESCENT, INITIATE A DESCENT, OR MAINTAIN ALTITUDE
2. STALL RECOVERY (ATS, UA)	USE THROTTLE (PROCEDURES)	TO INITIATE A RECOVERY WITH FULL POWER
3. EMERGENCY LANDINGS (HAE, LAE)	USE PROP/FLAPS/LANDING GEAR/CANOPY (PROCEDURES) AND DETERMINE LANDING AREA/WIND DIRECTION (ANTICIPATION)	TO CLEAN-UP FOR GLIDE, SELECT LANDING AREA, PREPARE FOR LANDING
4. CLEARING/BREAK TURNS (ELP, POS, PS)	ANTICIPATE AIRSPEED/POSITION (ANTICIPATION)	TO ROLL INTO AND OUT OF A TURN WHILE MAINTAINING A SPECIFIC AIRSPEED WITH NOSE ATTITUDE
5. SLOW FLIGHT CONFIGURATION (PROC, SF)	USE THROTTLE/PROP/FLAPS (PROCEDURES)	TO ENTER SLOW FLIGHT WHILE MAINTAINING ALTITUDE
6. STALL RECOVERY (NOSE LOW) (ACC SP, UA)	USE THROTTLE (PROCEDURES)	TO INITIATE A RECOVERY BY CLOSING POWER
7. COURSE RULE PERFORMANCE (HD, CR)	USE GEAR/CHECKLISTS (PROCEDURES), ANTICIPATE ALTITUDE AND DETERMINE LOCATION (ANTICIPATION)	TO ADHERE TO INFLIGHT OR FIELD DEPARTURE RULES

TABLE 3: CORRELATIONAL MATRIX AND MULTIPLE CORRELATIONS OF THE PREDICTOR FACTOR AND CRITERION PHASE VARIABLES

	CORRELATIONAL MATRIX (N=78)												MULTIPLE CORRELATIONS		
OVERALL PHASE	CONTINUOUS FACTORS**					DISCRETE FACTORS**							TOTAL	1ST	2ND
GRADES***	1	2	3	4	5	1	2	3	4	5	6	7	(N=78)	(N=39)	(N=39)
PRIMARY	.633	-	-	-	.502	.494	.322	.226	.282	.376	-	.319	.743	.701	.686
BASIC	.516	-	.249	-	.312	.423	.320	-	.269	.256	-	.327	.643	.627	.474
ADVANCED	.390	-	-	-	.266	.347	-	.230	-	.266	-	.265	.441	.527	.331**
AVERAGE CORRELATION	.513	0	.083	0	.360	.421	.214	.152	.184	.299	0	.304	-	-	-

* ALL CORRELATIONS ARE MINUS VALUES AND ARE SIGNIFICANT AT P .05

** THIS CORRELATION p .05, REST p .01

*** APPROXIMATE FLIGHT PHASE LENGTHS FOR AN AVERAGE TRAINING PERIOD OF 16 MONTHS: PRIMARY (0-4 MONTHS), BASIC (5-10 MONTHS), ADVANCED (11-16 MONTHS).

Proceedings of the [illegible]th Annual Meeting of the [illegible] Society
Los Angeles, CA [illegible]

THE VALIDITY OF TASK ANALYTIC INFORMATION TO HUMAN PERFORMANCE RESEARCH IN UNUSUAL ENVIRONMENTS

Richard H. Shannon
Naval Biodynamics Laboratory, New Orleans, LA 70189

ABSTRACT

Recurring naval student flight errors were collected using three types of task analytic methodologies: maneuver description questionnaires, critical incident technique of observed student problems and inflight maneuver rating forms. The results from these three efforts were highly correlated, which indicated that the task analytic methodology was valid; and therefore, would have utility in the future development of a performance battery for environmental research.

INTRODUCTION

Task analysis is a methodological tool which can be used: (1) to describe the functions performed by the human component, (2) to determine the relative position of each task on a certain dimension to the overall job, and (3) to specify the human capabilities necessary for criterion performance. The results of a valid task analysis can aid in the decision-making process concerning personnel selection, training and assessment as well as equipment design and test and evaluation. However, this procedure is not rigorous in the scientific sense, but is more heuristic, creative and innovative. There are guidelines as to how a task analysis should be conducted, but there are not any set rules that can reduce this technique to a routine one (Miller, 1971). Researchers in the literature do not always agree on method and definitions, and much of the work is fragmentary and inconsistent without utility or validity being established (Farina, 1969; Chambers, 1969). However, the literature can provide models upon which future analysts can develop their task descriptions. The purpose of this paper will be to determine the utility and validity of task analysis as well as to outline future intentions using the same methodology in other exotic environments besides aviation. To accomplish these goals, three related methodologies and their impact upon error isolation will be described under the headings of Maneuver Task Analysis, Critical Incident Technique of Grading Sheets, and Flight Rating Forms.

Maneuver Task Analysis

Prior to 1977 and the revision of the naval aviation training program, primary flight training in the T34 aircraft was divided into two stages, pre-solo and precision. Pre-solo consisted of dual instruction for the purpose of preparing the student pilot to fly solo while the precision stage was directed toward instruction in precision and acrobatic maneuvering. The pre-solo stage consisted of 12 dual instructional hops. By the ninth hop, all maneuvers had been introduced and the student began to prepare for his final check hop prior to solo. That is, after Hop 9, the student should have been able to execute all maneuvers without committing any major errors. Hop 12 was the safe-for-solo check flight, and if considered safe by the instructor, the student soloed on Hop 13. Beginning with Hop 4, the instructor rated a student's performance on the 21 maneuvers executed by assigning him either an above average, average, below average or unsatisfactory grade. The following maneuvers were selected for the task analytic methodology because they had a higher number of below average grades, and therefore, were considered to be very critical to student performance: standard field entry, high altitude emergency, precision spin, full flap approach/landing/takeoff, approach turn stall, and slow flight. These six maneuvers represented 63 percent of the total number of below average grades for the 70 students sampled. Three other items (basic airwork, headwork, procedures) on the instructor's rating form were not considered for analysis at this time but they had a higher number of below averages than the maneuvers selected. These items do not lend themselves to detailed descriptions since they are global skills encompassing all maneuvers.

Task maneuver descriptions were developed using the T-34 Primary Maneuver Description Handbook. Task inventory development was performed in gradually refined stages with the analysis proceeding from large units of information blocks to successively more detailed lower levels. A similar method was used by the author to describe naval flight officer and pilot functional tasks (Shannon, 1980a). The qualitative model utilized in this study followed a flight maneuver, sequential phase, sequential step, functional objective (task level) and task element hierarchial breakdown. Each block of information within this pyramid was considered to be a conceptual whole. Sequential phases were an arrangement of those sequential steps according to aircraft changes in fundamental flight attitude. These changes encompassed the initiation, maintenance and completion of either a climb, descent, turn, straight and level, or ground operation, which were considered separately or in combination (such as, climbing turn, descending turn). Sequential steps were distinct procedural operations that were performed serially along a time continuum. These steps contained functional objectives that occurred concurrently and were an arrangement of those task elements performing the same function. Task elements were considered to be parts of an information processing loop between the pilot, aircraft and the environment. This feedback loop contained the following components of display, sensory, cognition, motor and aircraft control. A task element statement contained an action verb with a definite object. Approximately 15 action verbs were classified according to a sensory, cognition, motor and communication taxonomy, and were utilized in these maneuver descriptions. The utility of a taxonomy and standardized action verbs is to allow the analyst to classify human activities, to have similar meaning and wording within task statements, and to perform a commonality analysis across maneuvers.

The initial maneuver descriptions utilizing this conceptual model and the T-34 Primary Maneuver Description handbook were checked for completeness and accuracy by seven flight intructors assigned to primary training. On the basis of their comments, suggestions and criticisms, revisions were made. Since the intent of this effort was to isolate recurring student problem areas during pre-solo training, these maneuver descriptions were given a questionnaire format with the following dichotomous scales: (1) Does this item represent a frequent error committed by the average student on all hops in primary training?, and (2) If the item is an error, is it critical? Critical error was described as one which affected a student's grade and was considered essential to the performance of the maneuver. These preliminary questionnaires were reviewed by four primary flight instructors in the same format to be used by the total sample. The purpose of this simulation was to gain insight into respondee attitudes toward the inventories. Their comments were then incorporated into a final version and distributed to 97 primary flight instructors, which was close to total participation by the squadron. The average length of these maneuver descriptions was 10 steps and 123 task elements.

The inventories were completed by 100% of the instructors. Upon their return, the data were coded for computer analysis. Percentages were assigned for each of the task elements according to the number of instructors who considered the element to be an error that was frequently performed by the average student and to be critical to his performance. In addition, percentages were also assigned at the functional objective and step levels according to the number of instructors responding to any of the elements within that objective or step. Since these two scales were considered to be independent, their percentages were multiplied together. This result was assumed to indicate the number of instructors who believed this item to be both frequent and critical. Stated another way, this percentage was considered to be the probability of either a task element, functional objective or sequential step containing errors which would result in a below average or unsatisfactory maneuver grade. A rank-ordering of these elements, objectives or steps by these probabilities of occurrence were an indication of the more serious flight problems of student pilots along both the time and type dimensions of maneuver performance.

A commonality analysis was performed. Each task element, functional objective and task activity were analyzed for similarity and were combined across the six maneuvers. Standardization of wording, meaning and function within the task element statements greatly assisted this effort. The result is the taxonomic outline in Table 1 which was to be used to classify student performance. Two operations, nine activities, and 35 functional objectives are shown, which conceptually depict flight performance as a joint function of continuous and discrete communications, scanning and controlling operations that occur in a four-dimensional inflight environment of pitch, roll, yaw and thrust. Discrete operations are defined as individually distinct movements or mediating responses elicited by environmental cues. Continuous operations contain those tasks involving multidimensional tracking responses to either contact cues outside the cockpit or flight instrument cues within the cockpit. An activity is one whose main function involves either a sensory, cognition, motor, or coordinated perceptual-motor task. A functional objective is defined as part of a maneuver having the same activity, performance goal, sensory cues, and task elements. A functional objective by step, which was the task or error level used in the maneuver descriptions, has the same definition as the preceeding statement. The difference between both categories is that one combines objectives over time and the other is limited to a sequential step or phase. The nine activities outlined in this paper are very similar to those used by other authors (Miller, 1971; Alluisi, 1967; Folley, 1964; Teichner, Whitehead, 1971; Kelley, Prosin, 1968). At the functional objective level, however, this congruence diverges. The following list of task activities and terminology from the literature will demonstrate this similarity: (1) monitor - scanning, detecting, identifying, monitoring, watchkeeping, sensing, searching; (2) procedural - short/long term memory, process, procedures- following, discrete motor, switching; (3) anticipation/planning - intellectual, decision-making, problem-solving, planning; (4) communication - communication; and (5) continuous pitch, roll, yaw, thrust, brakes - continuous perceptual-motor, tracking, action.

The importance of the taxonomic classification depicted in Table 1 is its utility as a tool and a structural model for future research into performance assessment, selection, training, and human engineering. Without a unifying system, Chambers (1969) believed it would be exceedingly difficult to achieve generalization of research results, communication between research and applied workers, application of research results to applied problems, and utilization of data acquired from one applied situation to another. This outline can now be used with other methodologies (critical incident technique, rating forms) in the determination of the size and scope of recurring student pilot error.

Critical Incident Technique of Grading Sheets

The critical incident technique (Flanagan, 1954) was applicable to the aims of this investigation, which was to determine the validity of the task analytic information and to establish the utility of the taxonomic classification of pilot behavior presented in Table 1. The incidents with which this technique deals are descriptions of directly observable human activities which are sufficiently complete in themselves to permit inferences to be made about the person performing the act. For the incident to be critical, it must describe segments of human behavior that are pertinent to a desired objective, such as student pilot error while performing flight maneuvers. This methodology outlines specific procedures whereby human error can be catalogued, described and analyzed.

On each flight after hop 4, the flight instructor had to rate a student's performance by listing on the front page either an above average, average, below average or unsatisfactory grade for each of the observed maneuvers. When a below average or unsatisfactory grade was assigned, the instructor had to justify this critical incident by specifying on the back page of the grading sheet the errors committed by that student. The flight records of 70 students were selected at random for analysis. Two criteria had to be met for final selection: (1) overall primary stage grade for each student had to be between 2.95 and 3.10, which was considered to be average performance, and (2) each of the students completed advanced flight training, and therefore, were designated naval aviators. A content analysis was performed on these written comments from the back pages of the grading sheets, and the errors were classified using the taxonomic model in Table 1.

The Naval Flight Student Reference Manual was developed from this analysis. This manual contains a detailed listing of the student errors in this study, which are classified according to maneuver, type (activity, objective, error) and time (step, phase). In addition, this information is divided into two parts: (1) flight maneuver section presenting 2029 errors by 660 below average/ unsatisfactory grades for 70 students, and (2) global item section outlining 1020 errors by 182 below average/unsatisfactory grades for 39 students. The 2029 errors listed for the 70 students in the flight maneuver section of the manual was used in the comparison between the task analytic and rating form data. From this data pool, typical procedural activity errors using the throttle control could be classified as to omission problems (no power addition, retraction or adjustment), sequence problems (power applied early/late, out of sequence wth another control, slow to adjust power), and commission problems (power reversals during retardation/ advancement, power setting high/low, rough throttle usage). In addition, the following errors were associated with the activity of continuous control in the pitch axis: (1) high, low or erratic control of nose attitude; (2) late, slow or no nose movement; (3) fast, slow or erratic control of airspeed; (4) gains, losses or erratic control of altitude; (5) overcontrol of or rough on flight stick; (6) late, not enough or no trim; (7) poor or no overall/external/internal scan pattern; (8) slow to see or does not see errors; and (9) slow to make or performs poor error corrections.

Flight Rating Forms

Flight maneuver rating forms were constructed on the basis of the same conceptual model utilized in the task analytic questionnaires and the critical incident analysis of the grading sheets. These forms were used to evaluate the performance of 21 students on the four hops immediately preceding the pre-solo check flight (hop 12). It was assumed that student performance was somewhat stable during this period, and therefore, a meaningful analysis could be conducted. In all, the following ten maneuvers were selected from the total 21 potential items (minus the three global items), which represented 85 percent of the total number of below averages assigned to the 70 students: slow flight, power-off stall, precision spin, approach turn stall, high altitude emergency, low altitude emergency, emergency landing practice, and full flap approach/ landing/ takeoff.

The rating forms were designed to be readable, easy to grade and small enough to be placed upon the pilot's kneeboard. Sequential phases and task activities were outlined, respectively, along the columns and rows. Functional objectives were abbreviated and placed within the matrix of phases and activities. A decision was made to use only one common functional objective by phase rather than having each step specified. The loss of information was considered to be small with a concurrent increase in the efficiency and effectiveness of data collection. To ensure user acceptance, two flight instructors evaluated the forms during a preliminary trial period. After this evaluation, comments and criticisms were noted and minor changes to the forms were incorporated. The rating checklists were then distributed to eight primary flight instructors, who were required to simply check each error when it occurred during maneuver performance. In addition, the instructors were requested to assign the usual grades of above average, average, below average, or unsatisfactory. In this way, the frequency and criticalness of the errors could be assessed.

The analysis of the performance by 21 students over the ten maneuvers indicated that there were 683 errors for 132 below averages, 899 errors for 255 averages, and 117 errors for 82 above averages. These results demonstrate an approximate linear relationship between grades and average number of errors with 5.2 and 1.4 errors, respectively, being assigned to below average and above average performance. The total number of below average errors were used in the comparison of data with the other two methodologies. An examination of the average number of errors by grade and task activity indicated that the most frequent and critical errors are pitch control, roll control, anticipation/planning and procedures. The average errors for each of the three grades over the eight task activities are similarly distributed as tested by a chi-square goodness of fit test (^{2} = .505, df = 16, nonsignificant). Two important conclusions can be derived from this analysis: (1) errors occur even when performance is considered proficient, and (2) the difference between the various levels of performance is mainly in the number of errors observed within specific task activities.

When the rating forms were initially developed and the study designed, the assumption was that student performance was relatively stable or reliable during the four flights immediately preceding the pre-solo check flight. To test this hypothesis, the number of maneuver errors on the most critical and frequent task activities of pitch, roll, anticipation/planning and procedures were compared between two flights. Fifteen students had been observed on two flights over the ten maneuvers. In other words, a fly-fly reliability was determined by comparing the performance of 15 students on four task activities (60 cases) on two consecutive days. The result was a correlation of .725 which meets the criterion for stability as established by the PETER battery development program. Another study (Wilcoxon, et al, 1952) which attempted to determine naval student flight reliability, resulted in a rather low correlation of .31. This poor result could possibly be explained by the project's emphasis upon specific error within each maneuver. For example, a student could demonstrate the same error or error type on different maneuvers over both days. This occurrence would then lower the reliability. If commonality of error across maneuvers had been considered in the Wilcoxon study, it is quite possible that its resultant reliability would have approximated the result in the present investigation.

Validity of Task Analysis

Recurring inflight errors by student naval aviators during primary training in the T-34 aircraft were isolated using three related methodologies: maneuver task analytic questionnaires, content analysis of student grading sheets, and inflight checklists of maneuver performance. The three efforts were independent in that there was a different instructor and student sample in each case. However, the results were comparable because each thrust had the same taxonomic model as depicted in Table 1. For this reason, each method is considered to be a task analytical procedure.

Concurrent validities of the three student error assessment techniques were determined through correlation of each method's results by sequential steps, functional objectives and functional objectives by steps within each maneuver. These correlations are shown in Table 2. The number of cases compared were dependent upon the number of maneuvers and the levels of analysis. The questionnaires collected data on six maneuvers. Functional objectives are the 35 listed in Table 1, and functional objectives by steps are these same 35 items at a more precise level, that of each procedural step. These results indicate a higher relationship between the rating forms/content of grading sheets than either comparison of questionnaires with the other two methods. All validity estimates are highly significant. In addition, the correlations remain very similar across the three levels of analysis within each of the comparisons. Two conclusions can be stated based upon these estimates: (1) flight instructors are consistent in their determination of student errors, and (2) they are more consistent when they are rating observed behavior rather than questionnaire material.

DISCUSSION

The task analytic procedure as outlined in this paper has demonstrated a high level of both content and concurrent validity as well as high reliabilities for both students and instructors. The utility of these results can be summarized: (1) a valid taxonomic model has been developed, (2) valid maneuver task descriptions have been outlined, (3) a clearer and more meaningful picture of student pilot behavior has emerged, (4) valid assessment of student pilots can be accomplished during the stable and reliable period of the three or four hops preceding the pre-solo check flight, (5) inflight maneuver ratings were developed which did elicit valid and reliable data and had flight instructor acceptance, and (6) the information determined from this study could have an impact upon naval student selection, training and assessment.

Presently, the Naval Biodynamics Laboratory (NBL) in New Orleans is attempting to develop a human performance battery (PETER, Performance Evaluation Tests for Environmental Research), which will be used in unusual and adverse environments. The major thrust of the program at this moment is to study various cognitive, perceptual and psychomotor tasks in order to determine their stability and sensitivity over repeated measurements in a laboratory setting. However, a task analysis of various U. S. Navy jobs and work stations must be conducted if the battery is to have applicability to human performance under actual and simulated conditions. The technique, which has been described in this paper, has the capability to not only define a particular job, but to isolate the critical components of that job. These elements in turn would provide the basis for a performance measurement system, which can be utilized in both the laboratory and the actual environment (Shannon, 1980b). In conclusion, the author believes that the methodologies and model discussed in this paper will aid in the development of functional inventories and in the assessment of performance in the oceanic environment, and therefore, will influence the growth of the NBL battery.

REFERENCES

Alluisi, E. A., Methodology in the Use of Synthetic Tasks to Assess Complex Performance. Human Factors, 1967, 9, 375-384.

Chambers, A., Development of a Taxonomy of Human Performance: A Heuristic Model for the Development of Classification Systems. AIR, Silver Spring, MD, 1969.

Farina, A. J., Jr., Development of a Taxonomy of Human Performance: A Review of Descriptive Schemes for Human Task Behavior. AIR, Washington, D.C., 1969.

Flanagan, J. C., The Critical Incident Technique. Psychological Bulletin, 1954, 51, 327-358.

Kelley, C. R., Prosin, D. J., Adaptive Performance Measurement. Dunlap and Associates, Inc., Santa Monica, CA, 1968.

TABLE 1: TAXONOMIC OUTLINE USED TO CLASSIFY STUDENT ERRORS

A. Continuous Operations - tasks involving multi-dimensional tracking responses.
   1. Pitch Axis Control - aircraft control in nose up/down axis.
      maintain (a) altitude, (b) airspeed, (c) nose attitude, (d) stick pressure
   2. Roll Axis Control - aircraft control in the wing up/down axis.
      maintain (a) angle of break, (b) distance, (c) heading, (d) rate of descent, (e) stick pressure
   3. Yaw Axis Control - aircraft control in the nose left/right axis.
      maintain (a) balanced flight, (b) heading, (c) rudder pressure
   4. Thrust Axis Control - aircraft control of forward movement.
      maintain (a) rate of descent, (b) throttle pressure
   5. Brake Control - aircraft control during ground operations by turning, stopping or changing speed.
      maintain (a) brake control

B. Discrete Operations - tasks involving individually distinct movements or mediating responses.
   1. Procedural - control of aircraft subsystems by not omitting, reordering or improperly performing necessary sequential steps.
      use (a) flaps, (b) landing gear, (c) fuel switches/lever, (d) canopy, (e) throttle, (f) prop, (g) battery/magnetoes, (h) equipment, (i) checklists
   2. Anticipation/Planning - tasks involving judgment, planning and "being-ahead" of aircraft.
      anticipate (a) aircraft, (b) position, (c) altitude, (d) airspeed,
      determine (e) wind direction, (f) landing site, (g) location
   3. Communication - tasks involving transfer of information from one source to another.
      communicate (a) verbally, (b) visually
   4. Monitor - searching and scanning inside/outside the cockpit for aircraft safety and maintenance of flight.
      scan (a) for aircraft/obstructions
      (b) temperature/pressure instruments

Miller, R. B., Development of a Taxonomy of Human Performance: a User-Oriented Approach, AIR, Silver Spring, MD, 1971.

Shannon, R. H., The Utility of Task Analytic Techniques to Research in Unusual Environments. Proceedings Aerospace Medical Association, 1980a.

Shannon, R. H., Task Analytic Approach to Human Performance Battery Development. Proceedings Human Factors Society, 1980b.

Teichner, W. H., Whitehead, J. Development of a Taxonomy of Human Performance: Evaluation of a Task Classification System for Generalizing Research Findings From a Data Base, AIR, Silver Spring, MD, 1971.

Wilcoxon, A. C., Johnson, W., and Golan, D. L. The Development and Tryout of Objective Check Flights in Pre-solo and Basic Instrument Stages of Naval Air Training, Naval School of Aviation Medicine, NAS, Pensacola, FL, 1952.

TABLE 2: CORRELATIONS OF STUDENT ERRORS BY COMPARING THE THREE TASK ANALYTICAL METHODOLOGIES AT THREE LEVELS OF ANALYSIS

	ANALYTIC METHODOLOGY*** (CASES IN PARENTHESIS)		
ANALYSIS LEVELS	(1)	(2)	(3)
STEPS	.578 (57)*	.527 (51)*	.830 (51)* .834 (89)**
OBJECTIVES	.594 (89)*	.585 (90)*	.821 (90)* .902 (157)**
OBJECTIVES BY STEPS	.556 (210)*	.468 (210)*	.730 (210)* .736 (362)**

* Comparison of data for six maneuvers
** Comparison of data for ten maneuvers
*** All correlations beyond .0001 significance level

(1) Questionnaires/Content Analysis
(2) Questionnaires/Rating Forms
(3) Content Analysis/Rating Forms

Presented at the 28th International Congress of Aviation and Space Medicine
Montreal, Canada, 8 - 11 September 1980

Performance Evaluation Tests for Environmental Research (PETER) using Task Analysis

Richard H. Shannon, Ph.D.
Naval Biodynamics Laboratory, New Orleans, LA 70189

INTRODUCTION

Task analysis is a methodological tool which can be used: (1) to describe the functions performed by the human component, (2) to determine the relative position of each task on a certain dimension to the overall job, and (3) to specify the human capabilities necessary for criterion performance. The results of a valid task analysis can aid in the decision-making process concerning personnel selection, training and assessment as well as equipment design and test and evaluation. However, this procedure is not rigorous in the scientific sense, but is more heuristic, creative and innovative. There are guidelines as to how a task analysis should be conducted, but there are not any set rules that can reduce this technique to a routine one (Miller, 1971). Researchers in the literature do not always agree on method and definitions, and much of the work is fragmentary and inconsistent without utility or validity being established (Farina, 1969; Chambers, 1969). However, the literature can provide models upon which future analysts can develop their task descriptions. For example, two different approaches to task analysis, hierarchial and time-line scenario, were used by the author to describe naval flight officer functions in eight navy aircraft (Shannon, 1980a), and naval student pilot behavior while performing six flight maneuvers (Shannon, 1980b).

Presently, the Naval Biodynamics Laboratory (NBDL) in New Orleans is attempting to develop a human performance battery (PETER, Performance Evaluation Tests for Environmental Research), which will be used to study behavior under unusual and adverse conditions (Kennedy, Bittner, Harbeson, 1980). To facilitate the accomplishment of this goal, a task analysis of various U. S. Navy jobs and work stations will be conducted. In general, the steps to be followed from inventory development to battery implementation will be: (1) Research and evaluate the existing literature on performance assessment; (2) Isolate the critical work stations onboard U. S. Navy ships which are important to its peacetime and combat missions; (3) Complete a task inventory of these critical work stations; (4) Develop a questionnaire in order to isolate the recurring critical elements pertaining to a particular work station; (5) Translate these elements into perceptual-psychomotor behavioral components; and (6) Design a perceptual-psychomotor battery to be used at NBDL which is compatible with the existing work of PETER.

Previously, a similar progression had been followed in the development of three task descriptions, which were used to evaluate naval student pilot training (Shannon, 1980a). Although this information was never used to aid in the development of a performance battery, most of the necessary work had been conducted. The purpose of this paper will be to discuss this prior effort because the outlined methodology, data collection and analyses will have a direct impact upon the future direction of the PETER program.

Method

Mechanics of Task Description

A statement of purpose is necessary from the beginning in order to give structure throughout the descriptive and analytical stages. This statement should contain the goals to be achieved and the personnel to be studied while using specified equipment components under certain environmental conditions. Another prerequisite for the analyst is the attainment of a detailed knowledge of the job, environment, and system to be investigated. This knowledge could be arrived at through expert opinion, on-site observations, research literature, interviews, open-ended questionnaires, existing task descriptions, procedural system handbooks, and pertinent human engineering standards documents.

Task inventory development is performed in gradually refined stages with the analysis proceeding from large units of information blocks to successively more detailed lower levels. The qualitative model of the three projects described in this paper followed a job, role, duty, task and element hierarchial breakdown. In this model, a job was defined by its roles, a role by its duties, a duty by its tasks, and a task by its elements. A task can also be defined as a unit of work which is directed toward the accomplishment of a goal. A task statement should contain an action verb with a definite objective, and should specify or imply the personnel, equipment and environmental conditions involved. Approximately 70 action verbs were classified, defined and utilized in these three studies based upon the following four-category task taxonomy: sensory/perception, cognition, communication and psychomotor. This taxonomy is in general agreement with the literature. The utility of a taxonomy, and standardized action verbs are to allow the analyst to classify human activities, to have similar meaning and wording within task statements, and to perform a commonality analysis across jobs. The remaining portions of this paper will now describe the three Navy projects for which the author developed task inventories.

Task Inventory Development

In 1973, the Chief of Naval Air Training (CNATRA) decided to upgrade, revise and evaluate pilot training. This program involved three pipelines (propeller, jet, helicopter), three phases (primary, basic, advanced) and eight types of aircraft. The author while stationed at NAMRL in Pensacola was the sole consultant during development of three task descriptions, one for each of the training pipelines. The descriptions listed all of the tasks an undergraduate pilot performed while he was in training. The inventories when completed were administered to all fleet replacement air groups (RAGs) in order to determine the following information on three five-point scales; (1) How frequently is the task actually performed at the RAG activity? (Frequency), (2) How important is it that students be trained in the task as undergraduates? (Criticality), and (3) How well does the training command presently train students in the task relative to RAG entry level requirements? (Training Adequacy).

In order to accomplish this task, it was necessary to have representatives from each of the pipelines and stages of undergraduate pilot training (UPT). These representatives were instructed in the "how of task writing" and were supervised throughout the development of the inventories by the author. To ensure a complete inventory of tasks, a matrix was used on each duty. This schematic contained a behavioral dimension of action verbs from the task taxonomy (Y axis) and a functional dimension of equipment or goals (X axis). The interior of the matrix was completed by the representative instructors with a simple "yes" or "no" answer to the question "Is this task done?".

The next step in the development of the inventories was the comparison of tasks throughout the various stages and phases of training. Primary phase tasks were compared with the tasks in the various stages of the Basic phase, which in turn were compared with tasks of the Advanced phase. At the conclusion of this step, there were three groups of tasks (jet, propeller and helicopter pipelines) which were standardized in meaning and wording. In addition, a commonality analysis was performed in order to determine the common tasks between pipelines and the specific tasks within pipelines. This effort would later allow for generalizations to be stated concerning the UPT program as well as cross-comparisons within and between fleet aviation communities.

The three inventories were then typed and administered to a limited number of training command instructors in the same questionnaire format to be used later in the replacement air groups. The purpose of this simulation was to gain insight into respondee attitudes toward the inventories. In addition, the training command instructors were asked to add or to consolidate tasks when necessary. These comments were then incorporated into a final version. The final three inventories were then prepared for replacement air group distribution. The average length of the inventories was 340 tasks, and the average number of common tasks across the three pipelines was 276 or 81 percent of the total. Each of the Marine and Navy replacement air groups on both the East and West Coasts were visited. The task analysis questionnaires were completed by most of the Instructor pilots on-board for an average completion rate of 88 percent.

Upon return from the field, all data were coded for computer analysis. Each of the tasks were analyzed by the quantitative data from each scale, and then the tasks were rank-ordered. An "underemphasized" task was considered to be one which was ranked 70% or higher on the "criticality" scale and 30% or lower on the "adequacy" scale. An "overemphasized" task had the reverse order. This filtering technique enabled the analysts to isolate those tasks which needed to be improved or added to training, and those tasks that could be reduced or omitted from training. The average number of "overemphasized" and "underemphasized" tasks across pipelines were, respectively, 4.3% and 7.2%.

Results and Discussion

A behavioral analysis was performed on the seventeen most underemphasized tasks in each of the three pipelines. The assumption is that these tasks (highly critical, inadequately trained) would form the most valid foundation of a performance battery to assess pilot behavior. In all, there were 37 different tasks which were studied. Individually, each flight task contains elements which can be considered as parts of an information processing loop between the pilot, aircraft and the environment. This feedback loop contains the following components of display, sensory, cognition, motor and aircraft control. This analysis dissected the 37 flight tasks into 194 task elements which were classified as to their behavior activities of scan, cognition, control and communication; flight operations of takeoff/landing and inflight; and flight conditions of normal and emergency/contingency. The results of this analysis can be seen in Table 1 which indicated that control, takeoff/landing and emergency/contingency elements had the highest number. The latter two classifications are more significant when one considers that most pilot functions are performed inflight under normal conditions. The result

TABLE 1

The Classification of the Underemphasized Task Elements by Activity, Operation and Condition

ACTIVITY		OPERATOR		CONDITION	
SCAN	54	TAKEOFF/LANDING	103	NORMAL	92
COGNITION	52	INFLIGHT	91	EMERGENCY/CONTINGENCY	102
CONTROL	74				
COMMUNICATION	14				

that motor functions are more important than nonmotor abilities (sensory, cognition) has been supported by other authors (Fleishman, 1966; Shannon, 1980c).

A content analysis of these 37 underemphasized tasks was performed. The tasks having the most commonality are listed below:

A. Scan

1. Scan inside the cockpit at pertinent instruments (altimeter, airspeed, needle/ball, navigation, temperature, pressure, fuel, etc.) which will present necessary information for flight maintenance (25).

2. Scan outside the cockpit (ground references, other aircraft, flight and wing attitudes) at pertinent cues and/or obstructions for purposes of safety, navigation and flight information (29).

B. Cognition

1. Calculate/apply aerodynamic principles for proper aircraft performance (during all types of takeoffs/landings/ditchings, stalls/spins, emergencies, hover) (44).

2. Assess system effectiveness and mission capability of the aircraft (through judgment, composure, malfunction isolation, crosscheck of degradation) (8).

C. Communication

1. Communicate information pertaining to aircraft performance (with controlling agencies or crew during emergencies/clearance/approach/departure; with maintenance personnel while recording system malfunctions) (9).

2. Apply proper communication procedures (during clearance, approach/ departure, emergencies) (5).

D. Control

1. Control aircraft during stall or low lift conditions (departed flight, spin, high angle of attack, power settling, autorotation) (18).

2. Control aircraft during takeoff/landing emergencies (aborted takeoff, engine-out waveoff/landing, ditched/forced landing) (18).

3. Control aircraft during emergencies while troubleshooting/coping with the situation (6).

4. Control aircraft during normal approach and landing (line-up, crosswind approach/landing/rollout, high gross weight takeoff/landing, landing rollout on wet/icy runway) (29).

5. Control aircraft during hover in response to crew chief directions (3).

The results of this analysis indicate that a battery which was designed to assess aviator performance should contain visual search/scan, problem-solving, continuous leg and two-handed tracking tasks, as well as simultaneous tests to evaluate an individual's time-sharing capability and judgment under stress. As demonstrated by these results, the technique, described in this paper has the capability to not only define a particular job, but to isolate the critical components of that job. These elements in turn can provide the basis for a performance measurement system, which can be utilized in both the laboratory and the actual environment (Shannon, 1980c). In conclusion, the author believes that the methodologies and model discussed in this paper will aid in the development of functional inventories in the oceanic environment, and therefore, will influence the growth of the NBDL battery.

References

Chambers, A., Development of a Taxonomy of Human Performance: A Heuristic Model for the Development of Classification Systems. AIR, Silver Spring, MD, 1969.

Farina, A. J., Jr., Development of a Taxonomy of Human Performance: A Review of Descriptive Schemes for Human Task Behavior. AIR, Washington, D. C., 1969.

Fleishman, E. A., Human Abilities and the Acquisition of Skill. In E. A. Bilodeau (ed.) Acquisition of Skill, New York: Academic Press, 1966, 147-167.

Kennedy, R. S., Bittner, A. C., Jr. & Harbeson, M. M. An engineering approach to the standardization of Performance Evaluation Tests for Environmental Research (PETER). Proceedings of the 11th Annual Conference of the Environmental Design and Research Association (EDRA) Charleston, SC. March 1980.

Miller, R. B., Development of a Taxonomy of Human Performance: a User-Oriented Approach. AIR, Silver Spring, MD, 1971.

Shannon, R. H., The Utility of Task Analytic Techniques to Research in Unusual Environments. Proceedings Aerospace Medical Association, 1980a.

Shannon, R. H., The Validity of Task Analytic Information to Human Performance Research in Unusual Environments. Proceedings Human Factors Society, 1980b.

Shannon, R. H., Task Analytic Approach to Human Performance Battery Development. Proceedings Human Factors Society, 1980c.

Proceedings of the 51st Meeting of the Aerospace Medical Association
Anaheim, CA 12 - 15 May 1980

THE UTILITY OF TASK ANALYTIC TECHNIQUES TO RESEARCH IN USUSUAL ENVIRONMENTS

RICHARD H. SHANNON
NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY DETACHMENT, NEW ORLEANS, LA 70189

Task analysis is a methodological tool which can be used: (1) to describe the functions performed by the human component in a system, (2) to determine the relative position of each task on a certain dimension to the overall job, and (3) to specify the human capabilities necessary for criterion performance. The results of a valid task analysis can aid in the decision-making process concerning personnel selection, training and assessment as well as equipment design and test and evaluation. However, this procedure is not rigorous in the scientific sense, but is more heuristic, creative and innovative. There are guidelines as to how a task analysis should be conducted, but there are not any set rules that can reduce this technique to a routine one. Researchers in the literature do not always agree on method and definitions, and much of the work is fragmentary and inconsistent without utility or validity being established. However, the literature can provide models upon which future analysts can develop their functional descriptions. The purpose of this paper will be to describe the task analytic approach used by this writer on three separate U. S. Navy contracts involving airborne functions and to outline future intentions using the same methodology in a different unusual environment.

Mechanics of Task Description

A statement of purpose is necessary from the beginning in order to give structure throughout the descriptive and analytical stages. This statement should contain the goals to be achieved and the personnel to be studied while using specified equipment components under certain environmental conditions. Another prerequisite for the analyst is the attainment of a detailed knowledge of the job, environment, and system to be investigated. This knowledge could be arrived at through expert opinion, on-site observations, research literature, interviews, open-ended questionnaires, existing task descriptions, procedural system handbooks, and pertinent human engineering standards documents.

Task inventory development is performed in gradually refined stages with the analysis proceeding from large units of information blocks to successively more detailed lower levels. The qualitative model of the three projects described in this paper followed a job, role, duty, task and element hierarchial breakdown. In this model, a job was defined by its roles, a role by its duties, a duty by its tasks, and a task by its elements. A task can also be defined as a unit of work which is directed toward the accomplishment of a goal. A task statement should contain an action verb with a definite objective, and should specify or imply the personnel, equipment and environmental conditions involved. Approximately 70 action verbs were classified, defined and utilized in these three studies based upon the following four-category task taxonomy: sensory/perception, cognition, communication and psychomotor. This taxonomy is in general agreement with the literature. The utility of a taxonomy and standardized action verbs is to allow the analyst to classify human activities, to have similar meaning and wording within task statements, and to perform a commonality analysis across jobs. The remaining portions of this paper will now describe the three Navy projects for which the author developed task inventories.

Naval Flight Officer Project (1, 2)

In 1971, the Chief of Naval Operations (CNO), recognized a need to revise the naval flight officer (NFO) training program and to determine future training equipment requirements and characteristics. In response to this need, the Naval Aerospace Medical Research Laboratory (NAMRL) in Pensacola, Florida was requested to conduct a series of investigations analyzing the operational functions of the NFO in the F14, F4, A6, P3C, P3B, E2B, EA6, RA5C. In all there were 13 positions in eight aircraft.

The F4 and F14 aircraft were selected as the initial positions to be described. This decision depended upon the existence of the F14 Readiness Air Group (VF124) and Aircrew Training Team whose members had operational experience in the F4 and were knowledgeable concerning the systems in the F14. These two inventories became the basis for the Naval Flight Officer Reference System, which was a pool of all NFO tasks in completed inventories by duty and role. For example, the F4/F14 tasks were the foundation for the A6 description, while the F4/F14/A6 item pool helped develop the P3C. Local technical advisors reviewed the NFO Reference System to select those task items which were identical and similar to their aircraft. An interview was then conducted involving an advisor's comments, and additional items were developed or existing tasks reworded. In this way, each inventory was standardized as to wording and meaning. This procedure enabled the performance at a later date of a commonality analysis, which indicated similarities and differences between NFO positions. Revisions, therefore, to the NFO training program could be recommended.

After the development of a preliminary inventory, selected members of a replacement air group (RAG) squadron for the subject aircraft reviewed the task content. Deletions, additions and rewording of task statements based upon these comments were incorporated into the final inventory which was then mailed to all the operational squadrons in the fleet. The average length of these functional description was 283 tasks. Two seven-point scales were used in the functional inventory to measure either the required proportion of time and effort (Part of Position) or the importance (Criticality) of each task/duty/role relative to the remaining tasks, duties, roles. When the completed questionnaires were returned, this information was coded, keypunched and analyzed. A "Part of Position" and "Criticality" mean, standard deviation, percent performing, and frequency distribution was given for each role, duty, and task. The average rate of analyzed to mailed inventories by NFO position was 67 percent.

Undergraduate Pilot Training Project(3)

In 1973, the Chief of Naval Air Training (NATRA) decided to upgrade, revise and evaluate pilot training. This program involved three pipelines (propeller, jet, helicopter), three phases (primary, basic, advanced) and eight types of aircraft. The author while stationed at NAMRL in Pensacola was the sole consultant during development of three task descriptions, one for each of the training pipelines. The descriptions listed all of the tasks an undergraduate pilot performed while he was in training. The inventories when completed were administered to all fleet replacement air groups (RAGs) in order to determine the following information on three five-point scales: (1) How frequently is the task actually performed at the RAG activity? (Frequency), (2) How important is it that students be trained in the task as undergraduates? (Criticality), and (3) How well does the training command presently train students in the task relative to RAG entry level requirements? (Training Adequacy).

In order to accomplish this task, it was necessary to have representatives from each of the pipelines and stages of undergraduate pilot training (UPT). These representatives were instructed in the "how of task writing" and were supervised throughout the development of the inventories by the author. To ensure a complete inventory of tasks, a matrix was used on each duty. This schematic contained a behavioral dimension of action verbs from the task taxonomy (Y axis) and a functional dimension of equipment or goals (x axis). The interior of the matrix was completed by the representative instructors with a simple "yes" or "no" answer to the question "Is this task done?".

The next step in the development of the inventories was the comparison of tasks throughout the various stages and phases of training. Primary phase tasks were compared with the tasks in the various stages of the Basic phase, which in turn were compared with tasks of the Advanced phase. At the conclusion of this step, there were three groups of tasks (jet, propeller and helicopter pipelines) which were standardized in meaning and wording. In addition, a commonality analysis was performed in order to determine the common tasks between pipelines and the specific tasks within pipelnes. This effort would later allow for generalizations to be stated concerning the UPT program as well as cross-comparisons within and between fleet aviation communities.

The three inventories were then typed and administered to a limited number of training command instructors in the same questionnaire format to be used later in the replacement air groups. The purpose of this simulation was to gain insight into respondee attitudes toward the inventories. In addition, the training command instructors were asked to add or consolidate tasks when necessary. These comments were then incorporated into a final version. The final three inventories were then prepared for replacement air group distribution. The average length of the inventories was [illegible] tasks, and the average number of common tasks across the three pipelines was [illegible]76 or 81 percent of the total. Each of the Marine and Navy replacement air groups on both the East and West Coasts were visited. The task analysis questionnaires were completed by most of the instructor pilots on-board for an average completion rate of 88 percent.

Upon return from the field, all data were coded for computer analysis. Each of the tasks were analyzed by the quantitative data from each scale, and then the tasks were rank-ordered. An "underemphasized" task was considered to be one which was ranked 70% or higher on either the "frequency" or "criticality" scales and [illegible]% or lower on the "adequacy" scale. An "overemphasized" task had the reverse order. This filtering technique enabled the analysts to isolate those tasks which needed to be improved or added to training, and those tasks that could be reduced or omitted from training. The average number of "overemphasized" and "underemphasized" tasks across pipelines were respectively [illegible]% and 7.2%.

F18 Task Description in Test and Evaluation(4)

The purpose of this effort was to explore the feasibility of the application of computer-aided techniques to the test and evaluation phase of the F18 aircraft system. Because this aircraft is single seated, the pilot would have more tasks than that of a two-man fighter such as the F4 or F14. During the testing of the F18, it was therefore considered imperative that the human engineering aspects of the system be scrutinized with extreme care and that rapid communication of deficiencies to decision-makers be implemented.

In response to this need, the Pacific Missile Test Center, Point Mugu, California funded the author in 1978 to "develop an F18 Task Taxonomy that depicts the activities required of an F18 pilot under VFR/IFR and day/night conditions and across all mission profiles." This tasking was given due to the author having "specialized knowledge not available currently at PACMISTESTCEN." The task inventory was to be a part of an ongoing project called Mission Operability Assessment Technique (MOAT), which is a tool for assessing man-system compatibility as related to mission success. The goal was to formulate a single interval scale for rating F18 tasks which contained elements of two separate scales - subsystem effectiveness and pilot workload. In addition, a portable IBM 5100 Computer would be used to aid the pilots in their ratings immediately after flight completion.

The pilot inventory was developed using the various McDonnell Aircraft Company publications on the F18, and the tasks contained in the "Controller of Aircraft" role of the jet undergraduate pilot task analysis and in the complete F4 Radar Intercept Officer's description. The hierarchial structure of the task description followed a role, duty, task outline as well as mission time-lines. The time-line classification included the phases of ground operations, takeoff/climbout, enroute, tactical, approach/landing. The role format was Controller of Aircraft, Weapons Manager and Support Systems Manager. The completed task description was checked by one of the F18 test pilots at Patuxent River, Maryland for completeness and accuracy. His comments were incorporated into the final version, which was then given to [illegible] Mugu. This inventory consisted of [illegible]88 common tasks, which expanded to approximately [illegible] tasks when the flight missions were considered.

Discussion

Although three different projects were discussed, the technique and model were similar for each study, and the developed inventories can be considered examples of a particular task analytic methodology. This paper also showed that there was a commonality of functions across jobs, missions and platforms. In general, the NFO and pilot descriptions had much content validity and utility. Both training programs were influenced by the results from the fleet. The FIS description of the MOAT project, on the other hand, still has to show its applicability. However, the pilots can revise the existing inventory while it is being used during the test and evaluation phase of development. This flexibility should eventually increase the inventory's validity and utility.

Presently, the NAMRL Detachment in New Orleans is attempting to develop a human performance battery (PETER, Performance Evaluation Tests for Environmental Research), which will be used in unusual and adverse environments. The major thrust of the program at this moment is to study various cognitive, perceptual and psychomotor tasks in order to determine their stability and sensitivity over repeated measurements in a laboratory setting. However, a task analysis of various U. S. Navy jobs and work stations must be conducted if the battery is to have applicability to human performance under actual and simulated conditions. This technique has the capability to not only describe a particular job, but to isolate the critical components of that job. These elements in turn would provide the basis for a performance measurement system, which can be utilized in both the laboratory and the actual environment. In conclusion, the author believes that the methodology and model discussed in this paper will aid in the development of functional inventories and in the assessment of performance in the oceanic environment, and therefore will influence the growth of the NAMRLD battery.

References

1. Naval Flight Officer Function Analysis, Vols I to VIII NAMRL, NAS, Pensacola, 1972.
2. Shelnutt, J. B. & Shannon, R. H. A Function Level Commonality Analysis of the F4/F14 NFO Positions. NAMRL, NAS, Pensacola, 1972.
3. Undergraduate Pilot Training Task Analysis: Phase I Report. Headquarters, CNATRA, NAS, Corpus Christi, TX. 1974.
4. Helm, W. R. & Donnell, M. L. System Operability: Concept and Measurement in Test and Evaluation. Proceedings of Human Factors Society, 1978.

Preprints of the 52nd Annual Scientific Meeting of the [illegible] Medical Association
San Antonio, Texas [illegible] May 1981

A COMPARISON OF TACTICAL NAVAL WORK STATIONS WITHIN THE AIR AND SEA ENVIRONMENTS

Richard H. Shannon & Robert C. Carter
NAVAL BIODYNAMICS LABORATORY, NEW ORLEANS, LA. 70189

INTRODUCTION

The Naval Biodynamics Laboratory is developing a human performance battery which will be used to study behavior under unusual and adverse conditions. To facilitate the accomplishment of this goal, a task analysis of various Navy jobs and work stations is being conducted. Prior efforts under this project which utilized this analytical tool can be found in Shannon (1980a,b,c,d). The present study extends this work by comparing two work stations which have similar tactical missions and whose tasks are performed in two different environments. The hypothesis used in this comparison is that there is a high commonality of human work functions across jobs having similar mission orientations even when there are differences in equipment, platform and environment. The two work stations analyzed are the tactical tasks of the Combat Information Center (CIC) of Navy ships and the Combat Information Control Officer (CICO) aboard the E2 aircraft. In addition, the statistical analyses of a preliminary battery to assess CIC performance will be discussed.

TASK INVENTORY DEVELOPMENT

Task inventory development is performed in gradually refined stages with the analysis proceeding from large units of information blocks to successively more detailed lower levels. The qualitative model of the two efforts described in this paper followed a job, role, duty, task, and element hierarchial breakdown. In this model, a job was defined by its roles, a role by its duties, a duty by its tasks, and a task by its elements. A task can also be defined as a unit of work which is directed toward the accomplishment of a goal.

A final inventory of 249 tasks pertaining to the Combat Information Control Officer (CICO) within the E2 aircraft was written. Content validity was established using subject-matter experts and existing task descriptions during inventory development. Two seven-point scales were used in the functional inventory to measure either the required proportion of time and effort (Part of Position) or the importance (Criticality) of each task, duty, and role. A sample of 35 people completed the task analytic questionnaires. Additional information concerning E2 task inventory development and analysis can be found in Shannon (1980a).

A preliminary task inventory concerning the Combat Information Center aboard Navy ships was developed using procedural systems handbooks and existing task descriptions. This information was [illegible] from handbooks outlining the NTDS and [illegible] ship systems, which presently are among the [illegible] components aboard [illegible]. The [illegible] task outlined in these manuals were compared with the tasks of the [illegible] the E2 aircraft. [illegible] was [illegible] inventory containing approximately [illegible] tasks performed [illegible] ship motion environment. The CICO inventory outlined tasks performed by two individuals in an airborne environment.

COMMONALITY ANALYSIS

The two work station descriptions were studied using the Position Analysis Questionnaire (PAQ) (McCormick, Jeanneret, and Mecham, 1972) in order to translate the specific tasks related to different environments, work stations, and equipment into common task elements. This questionnaire is composed of 194 elements which are separated into six divisions: information input, mental processes, work output, relationships with others, job context, and other job characteristics. The PAQ element scores were converted to 45 job factor scores by using factor loadings developed for 2200 jobs by McCormick et al. The 45 factors (dimensions) include 33 for the six divisions outlined above and 12 for the overall job.

Separate PAQ analyses were conducted by two different researchers, one for CIC and the other for CICO. Mecham, McCormick and Jeanneret (1977) suggest that agreement between PAQ analysts be indicated by a product-moment correlation of their 194 job element ratings. The correlation between both sets of ratings (CIC & CICO) was $r = .83$. Seven of the 194 elements were the major sources of disagreement between analysts, and without those elements the correlation was $r = .91$. These elements reflected differences in manning levels (e.g., CIC has more people, so there is more supervision) and environment (e.g., CICO includes airborne vibration, noise, and close quarters that are not characteristic of CIC). The two analysts' ratings were converted to factor scores on the 45 dimensions of the PAQ, and the resulting correlation between their factor scores was $r = .75$. If the six factors (job dimensions) related to manning levels and work environment were not considered, the correlation was $r = .87$. These results indicate that (a) the analysts are reliable, and (b) the CIC and CICO jobs are extremely similar despite differences of equipment, platform, and environment. Therefore, these correlations lend support to the hypothesis that jobs with similar mission orientations have high commonality of human work functions.

The following job dimensions (factors) derived from the PAQ were identified by [illegible] as being representative of the CIC and [illegible] jobs: [illegible] information [illegible] using [illegible] sources of information, and being aware of environmental conditions; [illegible] mental processing: [illegible] and processing information; [illegible] physical [illegible] [illegible] working [illegible] [illegible] [illegible] [illegible] [illegible] and performing clerical/related activities. The dimensions that are common to both the CIC and CICO work station typify jobs involving military tactical control. The rudiments of such jobs are command, control, communication, and coordination. It is hypothesized that jobs reflecting these characteristics will have similar job dimensions. In addition to the PAQ dimensions in common, CIC and CICO each have unique aspects. CICO was characterized by being in hazardous job situations, performing unstructured work, and working in an unpleasant/hazardous/demanding environment. The job dimensions unique to CIC were performing supervisory/coordination/related activities, exchanging job-related information, and supervising directing/estimating.

TEST SELECTION & EVALUATION

The assumption that various jobs can be composed of the same human work functions is a basis for the concept of job component validity (McCormick, 1979). The procedure for establishing validity includes: (a) identification of the work functions and their relative importance, (b) determination of human attributes associated with successful performance of the work functions, and (c) combination of the attribute requirements associated with each function into an estimate of the requirements for the entire job. If the job component validation is successful, then the human attributes and work functions acquire construct validity. Of course, a job component validity effort presumes that a taxonomy of work functions and a method for measuring all relevent human attributes are available. Both of these needs can be met through the use of the PAQ and the proper selection of psychological tests to measure human attributes (Carter, Kennedy, and Bittner, 1980).

What, then, are the human attributes required by these two jobs and how can they be measured? It was decided that those PAQ elements, which are rated as very important and are identified as a part of important job dimensions, represent the essential attributes of CIC and CICO. The 58 PAQ elements which qualified were compared with tests from the Kit of Cognitive Factors by Ekstrom, French, Harman, and Dermen (1976). One test was selected to represent each factor in the Kit, and then the tests were evaluated for their relationship to each of the 58 essential work elements. Seven tests were selected which had the highest number of PAQ elements. The emphasis of these tests was upon perceptual, reasoning, and numerical abilities.

These tests were administered as a pilot study to eight Navy enlisted volunteers once a day for 10 consecutive weekdays (Saturday and Sunday excluded). The order of the tests for all subjects was randomized between days but remained the same within days. Daily means, variances, and the interday (intertrial) correlations of each test were calculated. Two of the tests showed substantial linear increases of their mean scores across the 10-day experiment. Ninety-five percent of the variance of Hidden Figures daily means was linear, and 96 percent of the variance of the Maze Tracing means was linear. The slopes were .75 and 1.5 units per day, respectively. All other tests had less than 11 percent of the variance of their means attributable to a linear trend; the scores were not appreciably affected by practice. Similarly, the within-day variances of all seven tests were little affected by practice. In addition, intertrial correlations were studied in order to determine the reliability of measurement among the subjects from day to day. After all, a test is of little value if it measures a transient characteristic. Two of the tests, Building Memory, and Deciphering Language, did not measure a consistent factor across days. A one factor solution for each test showed that only 39 percent and 41 percent, respectively, of the variance was being explained. Another test, Maze Tracing, appeared to measure the same factor (spatial scanning) every time it was administered with 81 percent of the variance being explained across the 10 days. The remaining tests (Hidden Figures, Form Board, Addition, and Mathematics Aptitude) were of intermediate generalizability across days. They had 74, 71, 63, and 59 percent, respectively, of their variance across days explained using a single factor solution. In summary, several of the tests appeared to measure persistent attributes of the subjects and to be related to the CIC and CICO jobs. The more promising tests will be examined with larger groups of subjects. Kennedy, Carter, and Bittner (1980) have used a similar method to identify other tests that measure generalizable human attributes which may be useful for job component validation.

IMPLICATIONS

Identifying tests that are related to jobs is only the second phase of an effort to establish validity. The importance, or weight in the performance prediction equation, of each test is yet to be established. These weights could be obtained from the ratings of the importance of PAQ elements to each job, and from ratings of the extent to which each performance test represents the elements. The resulting weighted test battery could be used to predict performance on the job, or to study the effect of environmental stress on job performance. Furthermore, this information would provide a basis for synthesizing an analog to the job from the tests. By taking into account other important job attributes (like task sequencing, consequences of errors, rates of performance, and criticality of tasks), a simulated job could be constructed from the tests, just as complex molecules are constructed from a few simple elements in another domain of science. Job component validity is the first step in the evolution of a science of work behavior, and job analysis is the first step in job component validation.

(REFERENCES supplied by the authors upon request.)

END

DATE
FILMED

3-82

DTIC